

法政大学学術機関リポジトリ
HOSEI UNIVERSITY REPOSITORY

ウェアラブルデバイスから得た慣性データを用いた 性別・身長・体重分類

著者	河野 恵実
出版者	法政大学大学院情報科学研究科
雑誌名	法政大学大学院紀要．情報科学研究科編
巻	16
ページ	1-6
発行年	2021-03-24
URL	http://doi.org/10.15002/00023865

ウェアラブルデバイスから得た慣性データを用いた 性別・身長・体重分類 Classification of Gender, Height and Weight Using Inertial Data from Wearable Devices

河野 恵実

Megumi Kawano

法政大学大学院情報科学研究科情報科学専攻

E-mail: megumi.kawano.2w@stu.hosei.ac.jp

Abstract

In recent years, wearable devices become common and are used widely for life assistance. The sensors inside the wearables can be used not only for recognizing human daily activities but also classifying human attributes such as gender, height and weight. Almost all studies for human attribute classifications are based on a single device in a fixed position under a walking activity. Therefore, in this research, we study the classifications of gender, height and weight using six wearable devices that are placed in different body positions, namely chest, waist, left wrist, right wrist, thigh, and ankle. The twenty-nine participants were recruited and their inertial data in both walking and down-stare were obtained from the six wearables. We first perform classifications of gender, height, and weight based on each of the six devices using machine learning algorithms including SVM, Logistic Regression (LR), k-Nearest Neighbor (k-NN), and Random Forest (RF). It is found that the chest data is more effective for gender classification, wrist data for height classification, and chest or waist data for weight classification. We then perform decision fusion based on classified results from these devices in different activities. It was found that using a small number of sub-datasets can achieve a better classification result, and the fusion can increase the accuracy in classifications of gender, weight and height.

1. まえがき

IoT 時代といわれる昨今、ウェアラブルデバイスが注目を集めておりその普及率が増加している。アプリケーションの拡充や健康意識の高まり、業務効率化や人件削減を図る目的で今後の市場の拡大が見込まれている[1]。

そんな中ウェアラブルデバイスなどから得た慣性データをもとにした行動認識などの認識研究が盛んに行われている。その中の 1 つに人物属性認識がある。人物属性とは、性別・身長・体重・年齢などその人を表す生体的、身体的特徴の事を指す。これらの人物属性はソフトバイオメトリクスと呼ばれ、それだけでは一意に個人を判別する事は出来ないが、いくつかを組み合わせることで個人を判別する事が出来るのでないかといわれている。

Supervisor: Prof. Jianhua Ma

既存の人物属性認識では、データ取得行動が歩行のみである事が多い。日常生活では歩行動作以外にも多くの行動があるため、それでは識別できる環境が限られていて汎用性があるとは言えない。またデータ取得デバイスが単一であることが多く、複数デバイスでデータ取得を行った場合でも各デバイスで分類を行っていることが多い。行動認識の研究では、複数のデバイスを使用して分類を行った方が精度は高くなるといわれている。しかし、人物属性認識において複数デバイスからのデータを同時に使用した分類は数多くない。データを同時に使用することでそれぞれの特徴をうまく組み合わせることができ、より高精度に分類が行える可能性がある。

よって本研究では、ウェアラブルデバイスを用いて複数行動のデータを、体の 6 か所に固定したウェアラブルデバイスから取得し、それらのデータを同時に使用して性別・身長・体重分類を行うことを試みる。

ソフトバイオメトリクスを分類するメリットとして、その性別や年齢などのソフトバイオメトリクスに基づいたサービスの提供やデジタルサイネージの表示などがあげられる。よりの確に絞ったターゲットに広告を表示する事で、広告の効果を最大限にする事が出来る。また、その属性情報 1 つで個人が判別出来るバイオメトリクス情報と組み合わせることで、バイオメトリクス認証の際の処理時間を削減したり、精度の向上を図ることが出来る。認証の際にまず性別などのカテゴリでふるいにかけることで、精査対象の人数を減らし処理時間を短縮したり、少ない人数の中から選び出すことになるので精度も上がると考える。構成は以下の通りである。第 2 章で性別・身長・体重分類の関連研究について述べ、第 3 章で研究概要について述べる。第 4 章でデータ取得実験について、第 5 章でデータ処理について述べる。第 6 章で特徴量抽出について、第 7 章で使用した機械学習について述べる。第 8 章では分類結果の統合について、第 9 章では認識結果について述べる。そして最後に第 10 章でまとめと今後の課題について述べる。

2. 関連研究

本章では、関連研究について述べる。以下に、人物属性認識における関連研究を述べる。

Ankita らは、3 軸加速度とジャイロデータを用いて被験者の性別分類を行った[2]。被験者の前ポケットにスマー

トフォンを固定し、歩行時のデータを取得した。通常の速度、低速、高速の3パターン、各2回の歩行を行った。被験者あたり6個のデータを取得した。データの前処理として、正規化、線形補間を用いた100Hzへのリサンプリング、歩行サイクルの抽出を行い、ローカル特徴であるLBPとMLPを特徴量として抽出した。SVMとブーストラップで分類を行い、5分割交差検証で評価を行った。その結果、各歩行パターンで75%前後の精度を得られた。Sevisは、3軸加速度データを用いて被験者の性別、年齢、個人の認識を行った[3]。年齢はYoung, Adultの2値分類を行った。スマートフォンを用いて、40秒間のウォーキングとランニング時のデータの取得を行った。実世界の状態に近づけるため、実験時の服装、歩く場所、デバイスの固定箇所・向きは指定しなかった。抽出した特徴量は、平均、分散、標準偏差、最大値、最小値、合計、最大値と最小値の幅、尖度、歪度、二乗平均平方根、3軸の平均の二乗の合計の平方根の11種類31個で、WEKA機械学習ツールのバギング、J48(決定木)、Kstar(K平均法)、MLP(マルチレイヤーパーセプトロン)、Naïve Bayesを用いて分類を行った。性別・年齢分類どちらにおいても98%前後という高精度を得ることが出来た。

以下に結果の統合における関連研究について述べる。

Kumarらは、顔画像と指紋画像を用いて人物認識を行った[4]。特徴レベル、スコアレベル、決定レベルの3つのパターンでデータの統合を行い、結果の比較を行った。各統合方法の結果は、特徴レベルでは94%、スコアレベルでは86%、決定レベルでは75%となった。画像を用いた際は、特徴量を統合する事でより良い結果が得られることが分かった。Mangaiらは、パターン認識タスクにおける特徴量の融合と決定統合に使われる様々な手法のレビューを行った[5]。また、特徴量の融合と決定統合の両方の手法を組み合わせた新しいフレームワークの提案を行った。

これらの研究は、複数のデバイスを用いてデータの取得を行った場合でも、認識の前にデータの統合を行い、1度だけ認識を行って結果を出している研究がほとんどである。本研究では、各デバイスから得た慣性データごと一度認識を行い、それらの結果を統合する事で、人物属性認識を行う。複数のデバイスを用いてデータの取得を行った場合、データの段階で統合をおこなってしまうと、取得データに現れる各デバイスの特徴が最大限に生かれない可能性があると考ええる。デバイスごとに事前に分類を行いそれらの結果を統合する事で、デバイスごとの特徴が生かされ、より高精度な結果を得られるのではないかと考える。

3. 研究目的と概要

本章では、研究概要について述べる。図1に研究の流れを示す。まず初めに日常生活で行うと考えられる行動から慣性データの取得を行う。デバイスは胸部、腰部、手首、大腿部、足首の6か所に固定し、デバイスはスマートフォン、スマートウォッチを使用する。取得するデータは加速度データとジャイロデータの2種類である。推定する人物属性は、性別、身長、体重の3種類であり、

性別は男女の2値、身長はShort(S), Middle(M), High(H)の3値、体重はSmall(S), Middle(M), Large(L)の3値で推定を行う。データ取得後は、取得した生データに前処理を加え、特徴量を抽出する。抽出する特徴量は平均、分散、最大値、最小値、標準偏差である。抽出した特徴量を用いて機械学習による分類を行う。その後、機械学習によって得た結果の統合による分類も行う。それぞれの識別結果は、実際のクラスとどの程度あったかを認識率で評価する。

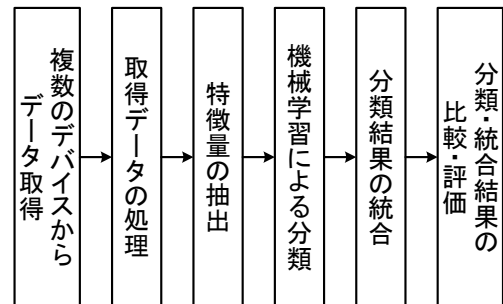


図1. 研究概要

本研究では、性別・身長・体重の分類の際に分類結果の統合を行う。分類結果の統合とはすべてのデータをいくつかのサブセットに分け、複数の分類器を用いて分類を行う。そこから得た結果を統合することで最終的な人物属性の決定を行う手法である。分類結果の統合を行うメリットとして、1)データをサブセットに分けることでデータの処理効率が上がる事、2)サブセットごとに適した分類器を使用する事で精度が向上する可能性がある事、3)すべてのテストデータに対して適切に機能する単一の分類器を見つける事は困難であるため、複数の分類器を組み合わせることでより優れた結果を得ることが出来る事が挙げられる。本研究では、6か所に固定した各デバイスのデータをサブセットに該当させ、それぞれのサブセットで分類器による分類を行い、それらから得た結果を統合して最良の結果を得られることを目指す。

4. データ取得

データ取得実験の流れについて述べる。各被験者の胸部、大腿部、両手首、足首、腰部の計6か所にスマートフォン、スマートウォッチを固定する。デバイス固定の様子を図3に示す。そのうえで、日常的に行う動作から各属性を分類する事を目的とし、被験者には日常的に行うと考えられる行動を行ってもらい、その行動中のデータを取得した。各行動3分を超えるように行った。それぞれの行動は大学構内で行ってもらい、被験者には20歳から24歳の男女計29名に参加してもらった。

識別する人物属性は、性別、身長、体重の3種類とする。性別は男女の2値、身長はSmall(~160cm), Middle(161~170cm), High(171cm~)の3値、体重はSmall(~54kg), Middle(55~65kg), Large(66kg~)の3値で分類を行う。身長と体重のクラス分けの値は22歳男女の平

均値から各クラスに該当する人数が均等になると思われる値に設定した．表 1 に各クラスの分類の閾値を示す．

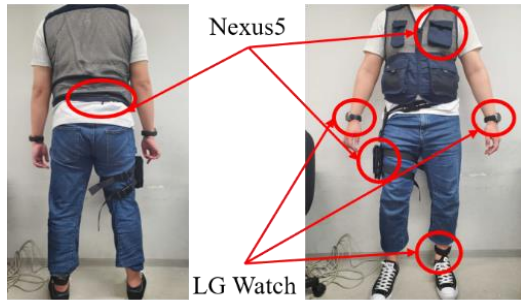


図 2. デバイス固定の様子

表 1. 分類属性の各閾値

Attribute	S	M	H/L
Height (cm)	～160	161～170	171～
Weight (kg)	～54	55～65	66～
Gender	Male	Female	

使用するスマートフォンは Nexus5 で，スマートフォンは G Watch Urban2 である．データ取得にあたり Android Studio でアプリケーションの作成を行った．センサの更新頻度の度に，タイムスタンプ，加速度データの X, Y, Z 軸の値，ジャイロデータの X, Y, Z 軸の値を csv ファイル形式で保存を行った．各デバイスのサンプリングレートは取得可能間隔の最短値に設定した．スマートフォンは 20ms でスマートウォッチは 10ms である．また，後に行うデータカットを行いやすくするため，データ取得を開始するデバイスの順を固定した．最初に胸部と右手首のデータ取得を開始し，その後腰部と左手首，大腿部と左足首の順にアプリを起動した．データ取得を終了する際は大腿部と左足首の取得を終了し，腰部と左手首，胸部と右手首の順にデータ取得を終了した．

5. データ処理

第 4 章で取得したデータには，実験に不必要なデータや，分類に悪影響を及ぼす細かい揺れによるノイズが含まれており，機械学習による識別の際に正しい結果が得られない可能性がある．実験に不必要なデータを取り除く為，データ処理を行う．まず初めにデータカットを行う．取得したままのデータの前後には必要な行動データとは関係の無いデータが含まれているためデータの前後をカットし，行動時のみのデータを切り出す．4 章で述べた通りの順でデータ取得を行っているため，左足首のデータを基準にカットを行う．続いて線形補間を行う．それぞれのデバイスで取得したデータはサンプリングレートや内部時刻の違いなど，様々な理由からデータのこうしん頻度，データ数が異なっている．特徴量抽出の際にデータ数が揃っている必要があり，認識の際には同じタイミングからのデータを使用しなければ認識精度が下がってしまう可能性があるため，データ更新のタイミングを統一し時間的ズレを解消する．

6. 特徴量データ

機械学習を行うにあたり，データの特徴をより分かりやすくするため，処理を行ったデータから特徴量の抽出を行う．本研究では，特徴量として平均，分散，標準偏差，最大値，最小値の 5 種類を使用する．それぞれの特徴量を加速度データ，ジャイロデータの X, Y, Z 軸から抽出した為，各部位ごと計 30 個の特徴量を抽出した．特徴量抽出の際はウィンドウサイズを 1000msec，オーバーラップを 500msec に設定した．現時点での特徴量データは，各特徴量データ間でスケールが異なっている．スケールが異なるままの特徴量を使用して，機械学習のモデルの学習を行うと，その後の分類で精度に悪影響を及ぼす可能性がある．よって学習の前段階でスケールを揃える必要があるため標準化を用いてスケーリングを行う．標準化はデータの平均を 0，標準偏差を 1 になるようにスケーリングを行う方法である．また，本研究は各デバイスからのデータで 30 種類の特徴量を算出し，データ段階で統合を行うと計 180 個の特徴量を扱うことになる．機械学習による分類において，そのすべての特徴量が識別に有効であるとは限らないため，機械学習による識別の前に特徴量選択を行う．本研究では RFE(Recursive Feature Elimination)を使用する．RFE は特徴に重みを割り当てる機械学習モデルを使用して，全特徴量を部分集合に分け学習・評価を多彩に性能が上がるか下がるかを確認，どの特徴量が識別に重要であるか判断していく．それを指定した特徴量数になるまで削除していく方法である．重み割り当ての機械学習に Random Forest を使用した．

7. 機械学習

分類問題において自力で閾値を決めて分類を行う事は難しい．よって，性別・身長・体重分類の際に機械学習を用いる．本研究では 4 つの機械学習を使用して精度を算出する．1 つ目は比較的認識精度が優れているとされている SVM である．教師あり機械学習法の 1 つで，未学習データに対しても比較的高い識別能力を持つとされている機械学習法である．分類の際にマージン最大化を用いることで汎化性を高め，高精度に識別が行える．2 つ目は最も簡単なアルゴリズムとされている k-NN である．クラス判別用の機械学習手法の一つで，既知の学習データをベクトル空間上にプロットし，新たに未知のデータが得られた時，そこから距離が近い順に任意の K 個を取得して，多数決でデータが属するクラスを決定する方法である．アルゴリズムが単純で分かりやすいとされている．3 つ目は RandomForest(RF)である．決定木を弱学習器として使用するアンサンブル学習である．今回は入力データの重複を許し複数のデータセットに分け，それぞれのデータセットを決定木で分類を行い，その予測結果を多数決する事で結果を得るブートストラップ法を用いる．4 つ目はロジスティック回帰(LR)である．教師あり学習の 1 つでベルヌーイ分布に従う統計的回帰モデルの 1 種である．ある事象が起きる確率の予測を行ったり，その確率をもとに分類を行う．確率 p の算出にはシグモイド関数を使用し，その確率が一定値以上であれば正事象のクラスに分類を行う．これらの機械学習による識別結果

の評価を行うのに交差検証を用いる。特に本研究では、データ数が少ないことと、各被験者の分類結果を確認したい目的から leave-one-out 交差検証を使用する。

8. 結果の統合

本研究では、複数のデバイスを使用してデータを取得し、取得した各デバイスのデータで性別・身長・体重の分類を行った後、その分類結果を統合して最終的な分類結果とする分類結果の統合を行う。統合の方法として、重み付き決定法と多数決法を用いる。

重み付き決定法(Weighted Sum, WS)は、統合前に分類を行った最小単位データセットの識別において、ラベルとそのラベルへの識別率を保持しておき、2つの積の合計を統合対象データ数で四捨五入した値を最終的な分類結果とする方法である。統合前の識別時、全ての各データセットにおいて、最も識別率の高かったラベルとその識別率だけではなく、全てのラベルとそのラベルへの識別率を保存しておく。そうすることによって、すべての要素を余すことなく使用する為、より正確な統合結果が得られると考えた。重み付き決定法のイメージ図を図3に示す。

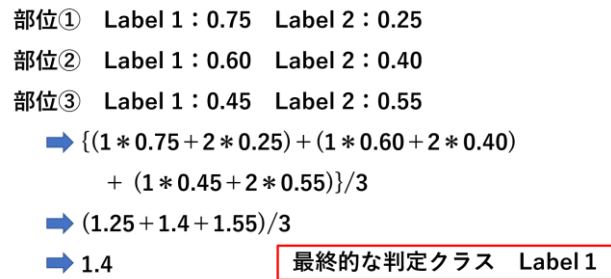


図 3. 重み付き決定法のイメージ図

多数決法(Majority Voting, MV)は、統合前に分類を行った最小単位データセットの識別結果のなかで、最も数の多い識別ラベルを最終的な分類結果とする方法である。最小単位での識別では、実際のラベルに関わらず、識別率が最も高かったラベルを分類の正解ラベルとする。統合の際には、その正解ラベルを用いて多数決を行い、最終的な正解ラベルを決定する。多数決の結果同数のラベルが存在した場合は、そのラベルへの識別率をすべて足し合わせた結果が大きい値となったものを正解ラベルとした。図4に多数決法のイメージ図を示す。



図 4. 多数決法のイメージ図

9. 性別・身長・体重推定の結果

本章では性別、身長、体重推定の結果についてそれぞれ述べる。分類の際には7章で述べた SVM, LR, k-NN, RF の4種類の機械学習法を用いた。

9.1. 性別分類

本節では性別分類の結果を示す。デバイスごとに分類を行った結果とその結果をもとに統合による分類結果を示す。性別分類の際の各クラス的人数は男性15人、女性14人である。

まず初めにデバイスごとの分類結果について述べる。取得した6つのデバイスそれぞれのデータのみを使用して分類を行った。歩行データからの分類結果を表2、階段下りデータからの分類結果を表3に示す。

表 2. 歩行データのデバイスごとの性別分類結果

Position	SVM	LR	k-NN	RF
Chest	74.78	75.26	68.35	56.01
Lumbar	51.70	53.42	62.12	65.24
Left Wrist	67.51	67.05	64.98	66.65
Right Wrist	67.65	67.39	64.52	64.97
Thigh	60.16	59.74	59.84	54.66
Left Ankle	74.55	73.96	66.56	68.46

表 3. 階段下りデータのデバイスごとの性別分類結果

Position	SVM	LR	k-NN	RF
Chest	62.65	61.91	60.19	58.23
Lumbar	50.81	50.30	56.78	53.98
Left Wrist	57.87	58.25	55.17	66.96
Right Wrist	52.46	52.58	50.46	57.82
Thigh	58.97	60.05	59.53	59.33
Left Ankle	60.14	59.54	52.78	59.60

行動を比較すると、歩行データを使用した方が高精度を得られた。階段下りは横の動きに加え、縦の動きも含まれるため、認識がしにくくなったと考えられる。部位ごとの精度を比較すると、胸部もしくは左足首からのデータを使用した際に比較的高い精度を得ることが出来た。これは歩行の際に現れる男女差が、胸部の動きと足の動きに顕著に出るためだと考える。男女は骨格の違いから足の出し方・重心移動の方法に違いがみられる[6]。そのため胸部と左足首からのデータを使用した際に高精度になったと考える。

続いて統合を行った結果について述べる。統合の際は、デバイスごとの分類結果をもとに、各デバイスでの正解率上位6, 5, 4, 3部位を使用した。表4に歩行データからの結果を統合した結果、表5に階段下りデータからの結果を統合した結果を示す。

統合方法ごとの精度を比較すると、性別の分類においてはあまり差が得られず、どちらの方法もある程度の精度を得ることが出来た。各デバイスでの分類で全て60%以上の精度が得られたため、6つ全てではなくても多くのデバイスで正しく分類が出来たのではないかと考える。

使用するデバイス数は少ない方が精度は高くなった．各デバイスの分類結果の精度が高い順に追加していったため，デバイスが増えていくと誤分類に影響されてしまったのではないかと考える．

表 4. 歩行データの統合による性別分類結果

	SVM		LR		k-NN		RF	
	WS	MV	WS	MV	WS	MV	WS	MV
6	78.57	82.14	85.71	82.14	82.14	85.71	78.57	78.57
5	85.71	85.71	92.86	85.71	82.14	82.14	75.0	78.57
4	89.29	89.29	89.29	89.29	89.29	89.29	78.57	82.14
3	92.86	92.86	92.86	92.86	78.57	82.14	75.0	82.14

表 5. 階段下りデータの統合による性別分類結果

	SVM		LR		k-NN		RF	
	WS	MV	WS	MV	WS	MV	WS	MV
6	72.41	72.41	68.97	68.97	72.41	68.97	68.97	68.97
5	75.86	75.86	75.86	72.41	72.41	72.41	68.97	72.41
4	79.31	79.31	82.76	79.31	72.41	72.41	79.31	79.31
3	68.97	65.52	68.97	68.97	68.97	72.41	75.00	78.57

9.2. 身長分類

本節では身長分類の結果を示す．身長分類の際の各クラスの人数は Short 9 人，Middle 11 人，High 8 人である．まず初めにデバイスごとの分類結果について述べる．歩行データからの分類結果を表 6，階段下りデータからの分類結果を表 7 に示す．

表 6. 歩行データのデバイスごとの身長分類結果

Position	SVM	LR	k-NN	RF
Chest	54.57	52.73	52.24	56.73
Lumbar	55.41	53.42	54.11	54.11
Left Wrist	57.93	57.82	43.87	58.58
Right Wrist	41.17	37.50	34.24	34.79
Thigh	41.72	42.03	41.32	40.68
Left Ankle	46.56	44.37	44.83	43.06

表 7. 階段下りデータのデバイスごとの身長分類結果

Position	SVM	LR	k-NN	RF
Chest	42.17	39.80	45.45	39.42
Lumbar	36.01	38.69	41.01	32.30
Left Wrist	42.31	41.41	40.78	40.64
Right Wrist	29.05	36.01	34.40	22.18
Thigh	36.27	37.51	33.10	36.71
Left Ankle	38.17	35.11	35.40	31.74

行動ごとのデータを比較すると，性別分類と同様に歩行データからの分類の方が高精度を得られた．部位ごとの結果を比較すると，左足首からのデータを使用すると比較的高精度を得ることが出来た．山崎ら[7]によると身

長が高くなるほど歩幅が大きく，歩行速度も速くなる．それに伴い動作中の腕の振りが大きくなったり，腰の動きに違いがみられ，高精度が得られたのではないかと考えられる．

続いて統合を行った結果について述べる．統合の際は，デバイスごとの分類結果をもとに，各デバイスでの正解率上位 6，5，4，3 部位を使用した．表 8 に歩行データからの結果を統合した結果，表 9 に階段下りデータからの結果を統合した結果を示す．

表 8. 歩行データの統合による身長分類結果

	SVM		LR		k-NN		RF	
	WS	MV	WS	MV	WS	MV	WS	MV
6	53.57	57.14	50.0	50.00	60.71	64.29	60.71	50.00
5	50.00	50.00	60.71	57.14	60.71	67.86	60.71	57.14
4	57.14	53.57	57.14	53.57	60.71	67.86	60.71	60.71
3	50.00	53.57	57.14	53.57	53.57	67.86	60.71	60.71

表 9. 階段下りデータの統合による身長分類結果

	SVM		LR		k-NN		RF	
	WS	MV	WS	MV	WS	MV	WS	MV
6	41.38	44.38	41.38	31.03	41.38	58.62	34.62	34.62
5	41.38	41.38	41.38	31.03	41.38	65.52	34.62	34.62
4	41.38	34.48	41.38	31.03	44.83	62.07	42.31	42.31
3	41.38	41.38	41.38	48.28	55.17	58.62	38.46	46.15

身長分類での統合の結果は，性別の時と比較して全体的に精度が悪化した．これはデバイスごとの分類結果の精度自体があまりよくなかったことが影響していると考ええる．各デバイスの分類で正しく分類出来ていなかった部位が多かったり，正しく分類されていてもその精度自体が低く，分類結果の悪いに部位に分類や精度が引っ張られてしまったと考える．

9.3. 体重分類

本節では体重分類の結果を示す．体重分類の際の各クラスの人数は Small 12 人，Middle 9 人，Large 7 人である．まず初めにデバイスごとの分類結果について述べる．歩行データからの分類結果を表 10，階段下りデータからの分類結果を表 11 に示す．

表 10. 歩行データのデバイスごとの体重分類結果

Position	SVM	LR	k-NN	RF
Chest	57.52	56.61	58.82	57.45
Lumbar	45.74	47.83	46.25	42.60
Left Wrist	46.67	43.17	45.31	41.93
Right Wrist	52.66	53.29	44.19	55.43
Thigh	53.54	53.95	46.06	57.75
Left Ankle	46.43	48.39	50.59	52.75

表 11. 階段下りデータのデバイスごとの体重分類結果

Position	SVM	LR	k-NN	RF
Chest	55.23	49.13	54.02	43.31
Lumbar	42.70	56.23	42.15	54.07
Left Wrist	42.72	40.33	39.83	43.71
Right Wrist	54.10	50.39	46.40	49.86
Thigh	48.55	37.76	44.36	33.60
Left Ankle	35.05	27.90	39.99	37.55

部位ごとのデータを比較すると、胸部データか腰部データを使用した際に比較的高精度が得られる。川野ら[8]によると体重が増えるにつれて重心の移動可能範囲が減少する。そのため体の重心近くの腰部や、重心から距離のある胸部の動きに違いがみられ、胸部のデータを使用した際に高精度が得られたと考える。

続いて統合を行った結果について述べる。統合の際は、デバイスごとの分類結果をもとに、各デバイスでの正解率上位 6, 5, 4, 3 部位を使用した。表 12 に歩行データからの結果を統合した結果、表 13 に階段下りデータからの結果を統合した結果を示す。

表 12. 歩行データの統合による体重分類結果

	SVM		LR		k-NN		RF	
	WS	MV	WS	MV	WS	MV	WS	MV
6	53.57	46.43	64.29	57.14	67.86	60.71	67.86	57.14
5	53.57	42.86	60.71	60.71	67.86	57.14	64.29	57.14
4	53.57	42.86	60.71	57.14	64.29	57.14	60.71	57.14
3	53.57	42.86	67.86	60.71	64.29	64.29	67.86	57.14

表 13. 階段下りデータの統合による体重分類結果

	SVM		LR		k-NN		RF	
	WS	MV	WS	MV	WS	MV	WS	MV
6	55.17	48.26	43.48	65.21	48.48	44.83	39.13	60.87
5	55.17	62.07	47.83	69.57	48.48	62.07	39.13	69.57
4	55.17	62.07	56.52	69.57	41.38	65.52	47.83	65.22
3	58.62	62.07	60.87	60.87	48.48	65.52	39.13	73.91

体重分類での統合の結果は、身長の時と比較すると精度は向上したが、性別の時と比較すると精度が下がった。どの機械学習法でも同じような値の精度が得られた。これは各デバイスの分類の際に分類クラスに偏りが生じてしまった為、統合を行っても正解クラスに再分類されることが難しく精度が低くなってしまったと考える。

10. まとめと今後の課題

本研究は体の 6 カ所に固定したウェアラブルデバイスから、歩行、階段下りの際の加速度データとジャイロデータを取得し、機械学習による性別・身長・体重分類を行った。分類の際には各デバイスのデータで分類を行った結果の統合を行う事で最終的な分類結果をする分類結果の統合も行った。各デバイスの分類のまとめとしては、性別の分類には胸、足からのデータが、身長の分類には手首からのデータが、体重の分類には胸、腰からの

データが識別に有効であることが挙げられる。統合のまとめとしては、データ取得を行ったすべての部位を使用するよりも、識別率上位何か所かの部位を使用する事でより精度が向上する事、統合を行う事で処理速度は高速になり精度もある程度向上することが挙げられる。

今後の課題として以下の点が挙げられる。1 つ目に被験者数の増加である。本研究での総被験者数 29 名、各クラスの被験者は 10 名前後となった。より汎化性を高めるためにはもっと多くのデータを集める必要がある。そのため更なる被験者数の増加が必要である。2 つ目にデータ取得環境の検討である。本研究では歩行、階段下りからの分類を試みが、実際の日常生活にはさらに多くの行動からなるため行動をさらに増やす必要があると考える。またデバイスやデバイス固定方法も検討が必要だと考える。3 つ目に最適な特徴量の検討である。今回は平均、分散、最大値、最小値、標準偏差の計 5 種類を用いた。さらに特徴の種類を増やすことで分類精度が向上するのではないかと考えられる。4 つ目に身長・体重クラスの閾値の変更、分類可能域の増加である。一般的に分類を行うためさらに細かいクラス分けや、各属性さらに値の大きい人、小さい人のデータが必要だと考える。5 つ目に統合方法の検討である。本研究では統合方法として、重み付き平均法と多数決法の 2 種類を使用した。2 種類だけでは十分とは言えないため他の方法の模索、精度の比較検討が必要だと考える。

文 献

- [1] 総務省, 令和元年番 情報通信情報通信白書, <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r01/pdf/n1200000.pdf> (2021 年 1 月 30 日参照)
- [2] A. Jain and V. Kanhangad, "Investigating Gender Recognition in Smartphones using Accelerometer and Gyroscope Sensor Readings", International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT), 623-628, May 2016.
- [3] K. N. Sevis, "Biometrics for SmartPhones: Age Recognition, Gender Recognition and Identification", Master Thesis, Isranbul Sehir Unibersity, 2017.
- [4] V. Kumar and R. Srikanthaswamy, "Comparative Analysis of Distinct Fusion Levels in Multimodal Biometrics", International Journal of Computer Applications, 2015.
- [5] U. G. Mangai, S. Samanta, S. Das and P. R. Chowdhury, "A Survey of Decision Fusion and Feature Fusion Strategies for Pattern Classification", IETE Technical Review, Vol.7, Issue4, 293-307, Sep, 2014.
- [6] AKIKO TAMURA Official Site, ~EBL~男女の歩行の違い, <https://akikotamura.net/archives/5485> (2021 年 1 月 30 日参照)
- [7] 山崎嶋廣, 佐藤陽彦, "ヒトの歩行一步幅, 歩調, 速度, およびエネルギー代謝の観点から一", 人類誌, P.385-P.401, 1990.
- [8] 川野常夫, 福井裕, 関子貴也, 吉田卓司, "体重増加に伴う重心移動の限界の測定", 日本人間工学会大会講演集, 2007.